AD-A215 238

DTIC FILE COPY

INFORMATION CAPACITY OF DIMENSION-LIMITED CHANNELS

C.R. Baker

Department of Statistics
University of North Carolina
Chapel Hill, NC 27599, U.S.A.

LISS-37

May 1989

DTIC
ELECTE
DEC 04 1989
S D
E

Abstract

Average information capacity is determined for a class of communication channels containing additive noise. Gaussian noise processes and a large class of nonGaussian processes are included. The constraint on the transmitted signals is given in terms of an increasing family of finite-dimensional subspaces. The results apply to the classical discrete-time channel and to continuous-time channels with fixed signal duration.

Research supported by ONR Grant N00014-89-J-1175 and NSF Grant NCR-8713726.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Department of Statistics 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) University of North Carolina Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-89-J-1175
8c. ADDRESS (City, State and ZIP Code) Statistics & Probability Program Arlington, VA 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification) Information capacity of dimension-limited channels	
12. PERSONAL AUTHOR(S) C.R. Baker	
13a. TYPE OF REPORT TECHNICAL 13b. TIME COVERED FROM ______ TO ______	14. DATE OF REPORT (Yr., Mo., Day) May 1989 15. PAGE COUNT 20
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Information capacity, Communication channels.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Average information capacity is determined for a class of communication channels containing additive noise. Gaussian noise processes and a large class of nonGaussian processes are included. The constraint on the transmitted signals is given in terms of an increasing family of finite-dimensional subspaces. The results apply to the classical discrete-time channel and to continuous-time channels with fixed signal duration.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189	22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED

Introduction

Coding capacity of additive Gaussian channels is a fundamental problem in the Shannon information theory [1], [4]. However, concrete results on this problem have actually been obtained only for special cases involving stationary noise.

In this paper, capacity is obtained for dimension-limited Gaussian channels. The result also holds for a large and important (in applications) class of nonGaussian channels. Bounds on capacity are obtained for a larger class of nonGaussian channels. The results given are for information capacity, but in a form suitable to use for obtaining coding capacity; that application is given elsewhere [3].

It will be seen that the development entails extensive use of the spectral representation of a self-adjoint linear operator. This operator defines the relation between the energy constraint on transmitted signals and the channel noise covariance. In particular, the essential spectrum of this operator plays (fittingly!) an essential role in the development.

Problem Formulation

The channel is described by $Y = X + N$, where Y, X, and N represent measurable stochastic processes in an underlying probability space (Ω, β, ν), with paths a.s. in a real separable Hilbert space H. H has inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. N is represented by a probability μ_N; μ_N need not be countably additive. Thus, μ_N is defined on the cylinder sets of H. Associated with μ_N is a covariance operator R_N that is linear, bounded, self-adjoint, and non-negative:

$$\langle R_N u, v \rangle = \int_H \langle x, u \rangle \langle x, v \rangle d\mu_N(x).$$

For
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

We assume WLOG that μ_N is strictly positive with zero mean.

In order to obtain finite capacity, any set of constraints on the signal X must involve (explicitly or implicitly) a self-adjoint bounded linear operator R_W which is strictly positive and satisfies

$$R_N = R_W^{1/2}(I+S)R_W^{1/2}.$$

In this representation $(I+S)^{-1}$ exists and is bounded, but the self-adjoint operator S need not be bounded. See [2] for details.

In order to formulate the problem in a form consistent for application to coding capacity, the constraints will involve an increasing family of linear manifolds (H_n), $H_n \subset H_{n+1} \subset H$, $\dim[H_n] = n$. $\tilde{P}_n$ will be the projection operator in H with range equal to H_n. P_{W_n} will be the projection operator with range equal to range$[R_W^{1/2} P_n]$.

For each n, the constraint on the class of admissible transmitted signals x is given as follows:

1) $\tilde{P}_n x \in H_n$ a.e. $d\mu_X$
2) $E_{\mu_X} \|\tilde{P}_n x\|^2_{W,n} \leq nP$

where μ_X is the cylindrical probability describing the transmitted signal, $P < \infty$ is fixed, and $\|y\|^2_{W,n} = \|u\|^2$, where u is the unique element in H_n satisfying $(\tilde{P}_n R_W \tilde{P}_n)^{1/2} u = y$. For each n, this defines a channel of the form $Y^n = X^n + N^n$, where X^n represents a stochastic process with paths a.s. in H_n and described by a (zero-mean) probability μ_X^n, $\mu_X^n(A) = \mu_X \circ P_n^{-1}[A]$ for each Borel set in H_n. N^n is defined by the zero mean probability $\mu_N^n = \mu_N \circ \tilde{P}_n^{-1}$.

The constraints (1) and (2) are equivalent to the following:

1') $\mu_X^n[H_n] = 1$

2') $E_{\mu_X^n} \|x\|^2_{W,n} \leq nP$

where μ_X^n is a probability. The equivalence of these two definitions follows from the fact that a cylindrical probability on a Hilbert space is uniquely defined by its projections on the finite-dimensional subspaces.

For any $n \geq 1$, let μ_{GN}^n be the zero-mean Gaussian probability on H_n with covariance operator $R_W^n = \tilde{P}_n R_N \tilde{P}_n$. The development here will make use of the relative entropy $H_{GN}^n(N)$, $n \geq 1$,

$$H_{GN}^n(N) \equiv \int_{H_n} \left[\log \frac{d\mu_N^n}{d\mu_{GN}^n}\right] d\mu_N^n.$$

We define $H_{GN}^n(N) = \infty$ if $d\mu_N^n/d\mu_{GN}^n$ does not exist. The most precise results (an exact value of the capacity) will be obtained for channels such that $\overline{\lim_n} \frac{1}{n} H_{GN}^n(N) = 0$. Of course, this includes all channels defined by Gaussian cylinder set measures; it also includes many nonGaussian channels. In fact, let $H_{GN}(N) \equiv \sup_n H_{GN}^n(N)$. Many important nonGaussian channels satisfy $H_{GN}(N) < \infty$. For example, if $\mu_N = \mu_V * \mu_X$, where μ_V is zero-mean Gaussian with covariance operator R_V and μ_X is zero mean with covariance operator $R_X = R_V^{1/2} T R_V^{1/2}$ for T trace-class, then $H_{GN}(N) < \infty$ regardless of the other characteristics of μ_X [3].

Two concrete examples of this setup are the classical discrete-time channel and the continuous-time channel with fixed (finite) transmission time for the code words. In the classical discrete-time channel, $H = \ell_2$, $H_n = \{x$ in ℓ_2: $x_i = 0$, $i > n\}$, and R_N is the noise covariance matrix. R_W is also defined by a covariance matrix; when the channel noise is stationary, then R_W is typically defined by a time-invariant linear filter h with Fourier transform $\hat{h}$. $|\hat{h}|^2$ then gives a spectral density which defines the covariance operator R_W. The present paper is limited to obtaining expressions for the

average information capacity, and related results, in the above framework. In [3], these results are applied to the classical discrete-time channel to obtain new results on coding capacity. Capacity of the stationary channel when $R_W = I$ has been known/assumed for many years; however, it is apparently only recently that a rigorous proof has been given for the channel with memory [5].

The other principal concrete example is for $H = L_2[0,T]$, $T < \infty$, with $H_n \equiv \text{span}\{e_1,\ldots,e_n\}$, where $\{e_k, k \geq 1\}$ is an infinite o.n. set. The application to such channels is described in [3], which also contains a number of specific results.

For each n, one thus has a channel in H_n, described by the joint probability μ^n_{XY}, $\mu^n_{XY}(A) = \mu^n_X \otimes \mu^n_N\{(x,y): x + y \in A\}$ for A a Borel set in $H_n \times H_n$. The mutual information is then

$$I(\mu^n_{XY}) = \sup_{\substack{k \leq K \\ K \geq 1}} \sum_{i=1}^{k} \mu^n_{XY}(A_i) \log\left[\frac{\mu^n_{XY}(A_i)}{\mu^n_X \otimes \mu^n_Y(A_i)}\right]$$

and the sup is over all Borel-measurable partitions $A_1, A_2, \ldots, A_k$ of $H_n \times H_n$. The capacity is $C^n_W(nP) = \sup I(\mu^n_{XY})$, where the sup is over all μ^n_X such that $E_{\mu^n_X} \|x\|^2_{W,n} \leq P$.

The channel's average information capacity is then $C^\infty_W(P) = \overline{\lim_{n\to\infty}} \frac{1}{n} C^n_W(nP)$.

Of course, coding capacity is the important operational parameter of the channel. If μ_N is Gaussian, then it can be shown [3] that coding capacity is equal to the average information capacity $C^\infty_W(P)$ for the class of channels considered here. In addition, if μ_N is not Gaussian, but satisfies $H^N_{GN} < \infty$, then $C^\infty_W(P)$ gives an upper bound on the coding capacity [3].

The program here is to give a general expression for $C_W^\infty(P)$, valid for all channels satisfying the above model. Such channels are described by a noise covariance operator R_N, a constraint operator R_W, an increasing family of finite-dimensional linear manifolds (H_n), and an "average signal-to-noise energy" limitation P. Equivalently, the channel is characterized by P, by S (the operator defining the relation between the noise covariance and the channel covariance), and by (P_{W_n}), a monotone increasing family of projection operators, P_{W_n} having range equal to range$(R_W^{1/2}\tilde{P}_n)$.

As part of the development, it will be seen that the results depend on the spectrum $\sigma(S)$ of S, but only upon the part belonging to the essential spectrum $\sigma_{ess}(S)$. This set, which is also the "limit points of the spectrum," consists of limit points of distinct eigenvalues, eigenvalues of infinite multiplicity, and points of the continuous spectrum [7].

Preliminary Results

The average capacity $C_W^\infty(P)$ is defined as $\overline{\lim_{n\to\infty}} \frac{1}{n} C_W^n(nP)$. $C_W^n(nP)$, the information capacity of the H_n channel, is given by the following well-known result.

Lemma 1 [2,5,6]. $\frac{1}{2} \sum_{i=1}^{N(n)} \log\left[\frac{B(n)+1}{\beta_i^n+1}\right] \le C_W^n(nP) \le \frac{1}{2} \sum_{i=1}^{N(n)} \log\left[\frac{B(n)+1}{\beta_i^n+1}\right] + H_{GN}^n(N)$ *where* $\beta_1^n \le \beta_2^n \le \ldots \le \beta_n^n$ *are the eigenvalues of* S_n, $N(n) = \sup\{i \le n: \beta_i^n \le B(n)\}$, *and* $B(n)$ *is defined by*

$$P = \frac{1}{n} \sum_{i=1}^{N(n)} (B(n) - \beta_i^n).$$

The next lemma gives several fundamental relations.

Lemma 2: Let $S: H \to H$ satisfy $R_N = R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}$, define $R_{W,n} = \tilde{P}_n R_W \tilde{P}_n$ and $R_{N,n} = \tilde{P}_n R_N \tilde{P}_n$, and let $S_n: H_n \to H_n$ satisfy $R_{N,n} = R_{W,n}^{\frac{1}{2}}(I_n + S_n)R_{W,n}^{\frac{1}{2}}$, for $n \geq 1$, where I_n is the identity in H_n. Then:

(1) $R_{W,n}^{\frac{1}{2}} = \tilde{P}_n R_W^{\frac{1}{2}} V_n^*$, where $V_n: H \to H_n$ is a partial isometry with initial set equal to range$(R_W^{\frac{1}{2}}\tilde{P}_n)$ and final set H_n.

(2) $S_n = V_n S V_n^*$, $n \geq 1$.

(3) Let P_{Wn} be the projection operator with range equal to $H_{Wn} \equiv$ range$(R_W^{\frac{1}{2}}P_n)$. The eigenvalues of $P_{Wn}SP_{Wn}$ (and their multiplicity) are the same as those of S_n.

Proof. (1) follows from results in [4]. (2) is obtained by equating the two definitions of $R_{N,n}$, yielding $\tilde{P}_n R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}\tilde{P}_n = R_{W,n}^{\frac{1}{2}}(I_n+S_n)R_W^{\frac{1}{2}}$, and then applying part (1). (3) follows from the fact that range$(P_{Wn}) = H_{Wn} = $ range(V_n^*). □

The family of partial isometries (V_n) plays a key role in the analysis of capacity. Consider $H = \ell_2$ and the classical discrete-time channel. Range$(R_W^{\frac{1}{2}}\tilde{P}_n)$, the range space of V_n^*, is then equal to the span of the first n columns of $R_W^{\frac{1}{2}}$. If $R_W = I$, this of course gives $V_n^* = V_n = \tilde{P}_n$, the projection onto $H_n = \{x: x_i = 0,\ i > n\}$. This is the simplest possible case, and enables one to immediately give the capacity of the stationary Gaussian channel as a corollary to the general expression for the capacity [3].

$\{e_i,\ i \geq 1\}$ will be used to denote any o.n. set in H such that $H_n =$ span$\{e_1, \ldots, e_n\}$. Similarly, $\{u_i,\ i \geq 1\}$ will denote any o.n. set such that $H_{Wn} = $ span$\{u_1, \ldots, u_n\}$. The partial isometry V_n is defined by $V_n u_i = e_i$, $i \leq n$. For fixed n, $\{v_i^n: i \leq n\}$ will denote o.n. eigenvectors of $P_{W_n}SP_{W_n}$ corresponding to the eigenvalues (β_i^n): $P_{W_n}SP_{W_n}v_i^n = \beta_i^n v_i^n$ for $i \leq n$.

Let $\{E_\lambda, \lambda \in \mathbb{R}\}$ be the left-continuous resolution of the identity [7] for S, so that for x in the domain $\mathfrak{D}(S)$ of S, $Sx = \int_{-1}^{\infty} \lambda dE_\lambda x$. For S bounded, this representation is the limit of a finite-limit-point approximation (FLPA) to S, (S^K), defined as follows. For S bounded, $1 \leq i \leq K$, let $-1 = a_0^K < a_1^K < \ldots < a_K^K$, where $1+a_{K-1}^K \leq \|I+S\| < 1+a_K^K$. Let Δ_i^K be the range of $E_{a_i^K} - E_{a_{i-1}^K}$, and set $E_{\Delta_i^K} \equiv E_{a_i^K} - E_{a_{i-1}^K}$. Let $S^K = \sum_{i=1}^{K} \theta_i^K E_{\Delta_i^K}$, where $\theta_i^K \in [a_{i-1}^K, a_i^K)$. Then $\|S^K - S\| \to 0$ as $K \to \infty$ and $\sup\{a_i^K - a_{i-1}^K : i \leq K\} \to 0$ [7]. If S is not bounded, then for any x in $\mathfrak{D}(S)$, $Sx = \lim_{M\to\infty} \sum_{n=0}^{M} \int_{n-1}^{n} \lambda dE_\lambda x$, where for each n, $\int_{n-1}^{n} \lambda dE_\lambda$ exists as the limit of a FLPA $(S^K[n])$ and $\|\int_{n-1}^{n} \lambda dE_\lambda - S^K[n]\| \to 0$ as $K \to \infty$ [7]. The operators (T_k), $T_k = \int_{k-1}^{k} \lambda dE_\lambda$, are pairwise orthogonal: $T_k : H \to H_k$, $H_k \perp H_j$ if $j \neq k$. Also, $T_k = \int_{k-1}^{k} \lambda dE_\lambda^n$, where E_λ^n is the restriction of E_λ to H_n. Thus, $H = \oplus_k H_k$, $E_\lambda^k : H \to H_k$, $Sx = T_k x$ whenever $x \in \mathfrak{D}(S) \cap H_k$.

If S is not bounded, then there exists a FLPA to S on $[-1, B]$ for any finite $B > 1$. That is, for any x in $\mathfrak{D}(S)$, $Sx = \lim_{M\to\infty} \sum_{k=0}^{M} \int_{k-1}^{k} \lambda dE_\lambda x$, where for each k, $\int_{k-1}^{k} \lambda dE_\lambda$ is the limit in the uniform operator topology of a FLPA. Thus, if B and B' are finite, then the bounded operators $E_B S$ and $E_{B'} S$ each have a FLPA, and the approximating sequences are consistent: if $B' > B$, then any FLPA (A^K) to $E_{B'} S$ provides a FLPA to $E_B S$, given by $(E_B A^K)$.

If (S^K) is a FLPA to (S), then $M_n^K(j)$ will denote the number of eigenvalues (repeated according to their multiplicity) of $P_{W_n}(S^K)P_{W_n}$ that lie in the interval $[a_{j-1}^K, a_j^K)$.

The integral representation of S will play a key role in the following analysis. For S bounded, $\|I+S\| \leq 1+\theta$, and f any continuous function on $(0, 1+\theta]$, it is noted that $[f(I+S)]x = \int_{-1}^{\theta+} f(1+\lambda) dE_\lambda x$ for all x in H [7].

Note that for any fixed n, the ordered eigenvalues of $(P_n S^K P_n)$ converge to those of $P_n S P_n$.

Lemma 3. *Let B be any positive real number, with f any continuous function on [0,B]. For all $n \geq 1$, define*

$$A_n = \frac{1}{n} \left| \sum_{j=1}^{n} I_{[0,B)}(1+\beta_j^n) f[1+\beta_j^n] - \lim_{K\to\infty} \sum_{j=1}^{K} I_{[0,B)}(1+\theta_j^K) f[1+\theta_j^K] M_n^K(j) \right|,$$

where $K \geq 1$ and $(I+S^K)$ is a FLPA to I+S. Then, $A_n \equiv 0$ for all $n \geq 1$, and the convergence is uniform for all $n \geq 1$.

Proof. First, assume that S is bounded and that $B > \|I+S\|$. Let $f(x) = x^p$ for any integer $p \geq 0$ and suppose that $1+\theta_K^K \leq M < \infty$, all $K \geq 1$. Then,

$$A_n = \frac{1}{n} \left| \sum_{i=1}^{n} (1+\beta_i^n)^p - \lim_{K\to\infty} \sum_{j=1}^{K} (1+\theta_j^K)^p M_n^K(j) \right|.$$

It is obvious that $A_n = 0$ if $p = 0$; thus, suppose $p \geq 1$.

$$A_n = \frac{1}{n} \left| \lim_{K\to\infty} \sum_{i=1}^{n} (1+\beta_i^{K,n})^p - \lim_{K\to\infty} \sum_{j=1}^{K} (1+\theta_j^K)^p M_n^K(j) \right|$$

where $(\beta_i^{K,n})$ are eigenvalues of $V_n S^K V_n^*$

$$= \frac{1}{n} \left| \lim_{K} \sum_{j=1}^{K} \left[\sum_{i:\beta_i^{K,n} \in [a_{j-1}^K, a_j^K)} (1+\beta_i^{K,n})^p - (1+\theta_j^K)^p M_n^K(j) \right] \right.$$

$$\leq \frac{1}{n} \lim_{K} \sum_{j=1}^{K} \left| \sum_{i:\beta_i^{K,n} \in [a_{j-1}^K, a_j^K)} (1+\beta_i^{K,n})^p - (1+\theta_j^K)^p M_n^K(j) \right|$$

$$\leq \frac{1}{n} \lim_{K} \sum_{j=1}^{K} \left[\sum_{i:\beta_i^{K,n} \in [a_{j-1}^K, a_j^K)} (1+\theta_j^K + a_j^K - a_{j-1}^K)^p - (1+\theta_j^K)^p M_n^K(j) \right]$$

$$\leq \frac{1}{n} \lim_{K} \sum_{j=1}^{K} M_n^K(j) \left[(1+\theta_j^K + a_j^K - a_{j-1}^K)^p - (1+\theta_K^K)^p \right]$$

$$\leq \frac{1}{n} \lim_{K} \sum_{j=1}^{K} M_n^K(j) \left[(1+\theta_K^K+\epsilon_K)^p - (1+\theta_K^K)^p \right]$$

where $\epsilon_K = \sup\{a_j^K - a_{j-1}^K,\ j \leq K\}$.

Now, $\sum_{j=1}^{K} M_n^K(j) = n$, every $K \geq 1$. Thus,

$$A_n \leq \lim_{K} \left[(1+\theta_K^K+\epsilon_K)^p - (1+\theta_K^K)^p \right] \leq \lim_{K} \left[(M+\epsilon_K)^p - M^p \right] = 0,$$

since the sequence (θ_K) is bounded and $\epsilon_K \to 0$. The same procedure obviously gives $A_n = 0$ when f is any polynomial. Note that the convergence is uniform for all $n \geq 1$.

Next, suppose that f is any continuous function on $[0,M]$, take $\epsilon > 0$, and let $Q = Q(\epsilon, f)$ be a polynomial such that $\sup_{0 \leq \lambda \leq M} |Q(\lambda) - f(\lambda)| < \epsilon$. Then

$$A_n \leq \frac{1}{n} \left| \sum_{i=1}^{n} Q(1+\beta_i^n) - \lim_{K \to \infty} \sum_{j=1}^{K} Q(1+\theta_j^K) M_n^K(j) \right| + \frac{1}{n} \left| n\epsilon + \epsilon \lim_{K \to \infty} \sum_{j=1}^{K} M_n^K(j) \right|$$

and the result follows as above.

To incorporate $B \leq \|I+S\|$, one makes an obvious modification in the above argument. Each FLPA (S^K) is chosen to contain B in its partition (a_i^K), and the sums are taken (for S^K) only over those j such that $1 + \theta_j^K < B$. Similarly, the (β_i^n) and $(\beta_i^{K,n})$ are summed only over those values $< B$.

For unbounded S, since n if fixed, for any $\epsilon > 0$ there exists $M(\epsilon)$ such that when $M > M(\epsilon)$, then $\| \sum_{n=0}^{M} \int_{n-1}^{n} \lambda \, dE_\lambda x - Sx\|^2 < \epsilon \|x\|^2$, for all x in H_{W_n}. That is, $\sum_{n=0}^{M} \int_{n-1}^{n} \lambda dE_\lambda$ converges to S in the strong (pointwise) operator topology on $\mathfrak{D}(S)$; however, restricted to the finite-dimensional subspace H_{W_n}, the convergence also takes place in the norm topology. One can thus apply the

preceding to the bounded operator $S_M = \sum_{n=0}^{M} \int_{n-1}^{n} \lambda dE_\lambda$, and then let $M \to \infty$ to obtain the result for S. □

Lemma 3 is a key result in determining the capacity. It will be applied to show that the capacity is completely determined (for a given set of constraints) by $(G_n, n \geq 1)$, where $G_n(\lambda) \equiv \frac{1}{n}$ [# eigenvalues of $S_n < \lambda$].

Several results that demonstrate the importance of the essential spectrum $\sigma_{ess}(S)$ to the channel capacity will now be obtained. Let $V[S,(H_n)]$ be the set of all γ in $\mathbb{R}$ such that for any K and any $\epsilon > 0$, there exists $n \geq K$ such that the number of elements in the sequence (β_i^n) satisfying $|\beta_i^n - \gamma| < \epsilon$ is $\geq K$. Note the importance of $V[S,(H_n)]$. For fixed λ, $\overline{\lim_n} \frac{1}{n}$ {# eigenvalues of $S_n < \lambda$] $= \overline{\lim_n} \frac{1}{n} \sum_{i=0}^{K-1}$ [# eigenvalues of S_n in $[a_i, a_{i+1})$] where $a_i < a_{i+1}$, $0 \leq i < K$, $a_0 = -1$, and $a_K = \lambda$. For $\gamma < \lambda$, let $a_i = (\gamma-\epsilon, \gamma+\epsilon)$. Then $\gamma \in V[S,(H_n)]$ is a necessary condition for $\overline{\lim} \frac{1}{n}$ [# eigenvalues of S_n in $(\gamma-\epsilon, \gamma+\epsilon)$] > 0 for every $\epsilon > 0$.

<u>Lemma 4</u>. *Let A be a K-dimensional subspace of H, with P_A the projection operator with range(P_A) = A. If $(K+1)\epsilon \leq 1$, then $\|P_A z_i\|^2 > 1 - \epsilon$ for at most K elements of the o.n. set $\{z_1, z_2, \ldots, z_{K+1}\}$.*

<u>Proof</u>. Let A be the span of the o.n. elements $g_1, \ldots, g_K$. Then $\sum_{i=1}^{K+1} \|P_A z_i\|^2 = \sum_{i=1}^{K+1} \sum_{k=1}^{K} \langle z_i, g_k\rangle^2 = \sum_{k=1}^{K} \sum_{i=1}^{K+1} \langle z_i, g_k\rangle^2 \leq K$. Thus, if $\|P_A z_i\|^2 > 1 - \epsilon$ for $i \leq K$, then $\|P_A z_{K+1}\|^2 < K - K(1-\epsilon) = K\epsilon \leq 1 - \epsilon$ if $\epsilon \leq \frac{1}{K+1}$. □

<u>Prop. 1</u>. *Suppose that $\{u_n, n \geq 1\}$ is a c.o.n. set for H. Let $\{z_n, n \geq 1\}$ be an o.n. sequence in H such that $\|(S-\gamma I)z_n\|^2 \to 0$. Then $\gamma \in V[S,(H_n)]$.*

Remark. The existence of an infinite o.n. sequence (z_n) such that $\|(S-\gamma I)z_n\|^2 \to 0$ is a necessary and sufficient condition for $\gamma \in \sigma_{ess}(S)$ [7]. Thus, Prop. 1 states that $\sigma_{ess}[S] \subset V[S,(H_n)]$.

Proof. Suppose that there exists $\epsilon > 0$ and $K \geq 1$ such that $\dim[\mathrm{span}\{v_i^n\colon |\beta_i^n-\gamma| < \epsilon\}] \leq K$ for all $n > 1$. By assumption, and using the fact that the monotone family of projection operators (P_{W_n}) converges in the strong topology to the identity,

$$0 = \lim_i \|(S-\gamma I)z_i\|^2 = \lim_i \lim_n \|(S-\gamma I)P_{W_n} z_i\|^2$$

$$\geq \lim_i \lim_n \|P_{W_n}(S-\gamma I)P_{W_n} z_i\|^2, \qquad \text{so that}$$

$$0 = \lim_i \lim_n \sum_{j=1}^{n} (\beta_j^n - \gamma)^2 \langle z_i, v_j^n\rangle^2.$$

Thus, for any $\delta > 0$ there exists $i_0(\delta)$ such that for all $i > i_0(\delta)$,

$$\delta > \lim_n \sum_{j=1}^{n} (\beta_j^n - \gamma)^2 \langle z_i, v_j^n\rangle^2.$$

Ordering $\{v_i^n, i \leq n\}$ such that $|\beta_i^n - \gamma| \geq \epsilon$ for $K+1 \leq i \leq n$, we have $\delta/\epsilon > \lim_n \sum_{j=K+1}^{n} \langle z_i, v_j^n\rangle^2$ for $i > i_0(\delta)$. For fixed δ, choose $k_1, k_2, \ldots, k_{K+1}$ such that $k_i > i_0(\delta)$ for $1 \leq i \leq K+1$. For any $\alpha > 0$, $\Delta > 0$, there exists $n_0(\alpha,\Delta)$ such that $n > n_0(\alpha,\Delta)$ implies that $\|P_{W_n} z_{k_i}\|^2 > 1 - \Delta$ and $\delta/\epsilon + \alpha > \sum_{j=K+1}^{n} \langle z_{k_i}, v_j^n\rangle^2$ for $1 \leq i \leq K+1$. For such an n, let A_n be the span of $\{v_1^n, \ldots, v_K^n\}$. Then for $i \leq K+1$, $\|P_{A_n} z_{k_i}\|^2 \geq \|P_{W_n} z_{k_i}\|^2 - (\delta/\epsilon + \alpha) > 1 - \Delta - (\delta/\epsilon + \alpha)$. Since Δ, δ, and α can be selected so that $[K+1](\Delta + \delta/\epsilon + \alpha) < 1$, one obtains a contradiction of Lemma 4. $\square$

Prop. 1 gives $V[S,(H_n)] \supset \sigma_{ess}(S)$, so long as $\{u_n, n \geq 1\}$ is complete. A partial converse is given by the following result.

<u>Prop. 2</u>. *Let θ_L and θ_U denote the smallest and largest points in $\sigma_{ess}(S)$. Then $V[S,(H_n)] \subset [\theta_L, \theta_U]$.*

<u>Proof</u>. It will first be shown that no point in $V[S, (H_n)]$ can exceed θ_U when S is bounded. γ is in $\sigma_{ess}(S)$ if and only if $\|(S-\gamma I)z_n\|^2 \to 0$ for some o.n. sequence (z_n), and this requires that $\langle Sz_n, z_n\rangle \to \gamma$. In order that $\gamma > \theta_U$ be in $V[S, (H_n)]$, it is necessary that for any $\epsilon > 0$ and $M < \infty$ there exists $n \geq M$ such that $|\beta^n_{i(\gamma)} - \gamma| < \epsilon$ for $i(\gamma) = 1,2,\ldots,M$. This requires that $|\langle Sz_i, z_i\rangle - \gamma| < \epsilon$ for at least M o.n. elements (z_i) in H. For $\delta > 0$, let $A(\theta_U, \delta)$ be the set of normalized x in $\mathfrak{D}(S)$ such that x is an eigenvector of S_n for some $n \geq 1$ and the corresponding eigenvalue is $\geq \theta_U + \delta$. Let M^δ be the maximal number of o.n. elements in $A(\theta_U, \delta)$. $M^\delta < \infty$ for every $\delta > 0$ implies that $V[S, (H_n)]$ contains no number greater than θ_U.

The operator $I - E_{\theta_U+\delta/2}$ is compact, since it must have finite-dimensional range space. Hence, if M^δ is not finite, then for any $n \geq 1$ there exists z_n in $A(\theta_U, \delta)$ with $\|z_n\| = 1$ and $\|(I - E_{\theta_U+\delta/2})z_n\| < \frac{1}{n}$. Let $\tau = \sup_{\|x\|=1} \langle Sx, x\rangle$. Then, for any $\epsilon > 0$,

$$1 + \theta_U + \delta < \langle (I+S)z_n, z_n\rangle = \int_{-1}^{\tau+\epsilon} (1+\lambda)d\|E_\lambda z_n\|^2$$

$$= \int_{-1}^{\theta_U+\delta/2} (1+\lambda)d\|E_\lambda z_n\|^2 + \int_{\theta_U+\delta/2}^{\tau+\epsilon} (1+\lambda)d\|E_\lambda z_n\|^2$$

$$\leq \int_{-1}^{\theta_U+\delta/2} (1+\lambda)d\|E_\lambda z_n\|^2 + (1+\tau+\epsilon)\|(I - E_{\theta_U+\delta/2})z_n\|^2$$

$$\leq \int_{-1}^{\theta_U+\delta/2} (1+\lambda)d\|E_\lambda z_n\|^2 + (1+\tau+\epsilon)/n$$

$$\leq (1+\theta_U+\delta/2) + (1+\tau+\epsilon)/n.$$

For sufficiently large n, this is a contradiction.

A similar approach shows that every number in $V[S, (H_n)]$ is $\geq \theta_L$, since $E_{\theta_L-\delta}$ has finite-dimensional range for every $\delta > 0$. □

The results of Prop. 1 and Prop. 2 lead naturally to several reasonable hypotheses: that $V[S,(H_n)] \subset \sigma_{ess}(S)$, that $V[S,(H_n)] = \sigma_{ess}(S)$ when $\{u_n, n \geq 1\}$ is complete, and that $\overline{\lim_n} \frac{1}{n}$ [# eigenvalues of $S_n < \lambda$] is independent of the choice of the c.o.n.s. $\{u_n, n \geq 1\}$. Unfortunately, all three of these desirable attributes are false, in general.

<u>Prop. 3</u>.

(1) *If $\{u_n, n \geq 1\}$ is not complete for H, then $V[S,(H_n)] \cap \sigma_{ess}(S) = \phi$ is possible.*

(2) *If $\{u_n, n \geq 1\}$ is complete for H, then:*

(a) *$V[S,(H_n)] \subset \sigma_{ess}(S)$ is not always true.*

(b) *Let $Q_\Delta(S,H_n)$ be the number of eigenvalues β_i^n of S_n such that $|\beta_i^n - x| \geq \Delta > 0$ for all x in $\sigma_{ess}(S)$. Then $\overline{\lim} \frac{1}{n} Q_\Delta(S,H_n)$ can be strictly positive.*

(c) *$\overline{\lim_n} \frac{1}{n}$ [# eigenvalues of $S_n > \lambda$] may not be independent of the choice of $\{u_n, n \geq 1\}$.*

<u>Proof</u>. Counterexamples will be constructed, each using the operator S defined as follows. Let $\{v_n, n \geq 1\}$ be any c.o.n. set in H, take $\beta > \alpha > 0$, and let

$$S = \alpha \sum_{\substack{n\geq 1 \\ n \text{ odd}}} v_n \otimes v_n + \beta \sum_{\substack{n\geq 2 \\ n \text{ even}}} v_n \otimes v_n .$$

To prove (1), set

$$u_n = \frac{v_{2n-1} + v_{2n}}{\sqrt{2}} \qquad n \geq 1.$$

Then $(P_{W_n} S P_{W_n}) u_i = \left[\frac{\alpha+\beta}{2}\right] u_i$, $i \leq n$, $n \geq 1$, while $\sigma_{ess}(S) = \{\alpha, \beta\}$.

To prove 2(a), choose $\{u_n, n > 1\}$ by

$$u_n = \frac{v_n + v_{n+(k+1)10^{k+1}-k10^k}}{\sqrt{2}} \qquad 2k10^k < n \leq (k+1)10^{k+1} + k10^k$$

$$= \frac{v_n - v_{n-(k+1)10^{k+1}+k10^k}}{\sqrt{2}} \qquad (k+1)10^{k+1} + k10^k < n \leq 2(k+1)10^{k+1}$$

for $k = 0,1,2,\ldots$ (β_i^n) then has the following composition:

$n = (2k)10^k$: $(\beta_i^n) = (\alpha, \beta$, each of multiplicity $k10^k)$

$n = (k+1)10^{k+1} + k10^k$: $(\beta_i^n) = (\alpha, \beta$, each of multiplicity $k10^k)$
$\cup$ $(\frac{\alpha+\beta}{2}$, each of multiplicity $(k+1)10^{k+1} - k10^k$.

For 2(b), note that if $\frac{\alpha+\beta}{2} < \lambda < \beta$, then the above choice of $\{u_n, n \geq 1\}$ gives $\overline{\lim_n} \frac{1}{n}$ [# eigenvalues of $S_n < \lambda] = \lim_{k\to\infty} \frac{(k+1)10^{k+1}}{(k+1)10^{k+1}+k10^k} = \frac{10}{11}$. Further, for $\Delta > (\frac{\beta}{2} - \frac{\alpha}{2})$, $\overline{\lim_n} \frac{1}{n} Q_\Delta(S, H_n) = \lim_{k\to\infty} \frac{(k+1)10^{k+1}-k10^k}{(k+1)10^{k+1}+k10^k} = \frac{9}{11}$.

Finally, 2(c) is shown by using $\{u_n, n \geq 1\}$ defined by $u_n = v_n$, $n \geq 1$, which yields (β_i^n) containing only α and β, each of multiplicity $n/2$ (n even) or α with multiplicity $\frac{n+1}{2}$, β with multiplicity $\frac{n-1}{2}$ (n odd). □

Average Information Capacity

The main result can now be obtained.

Theorem.

(1) $$\overline{\lim_{n\to\infty}}\ \tfrac{1}{2}\int_{-1}^{B_n} \log\left[\frac{B_n+1}{\lambda+1}\right]dF_n(\lambda) \le C_W^{\infty}(P)$$

$$\le \overline{\lim_{n\to\infty}}\left[\tfrac{1}{2}\int_{-1}^{B_n} \log\left[\frac{B_n+1}{\lambda+1}\right]dF_n(\lambda) + \frac{1}{n} H_{GN}^{n}(N)\right]$$

where B_n $(n \ge 1)$ *is defined by*

$$P = \int_{-1}^{B_n} (B_n-\lambda)dF_n(\lambda)$$

and

$$F_n(\lambda) = \frac{1}{n}\,[\#\ \text{eigenvalues of}\ S_n \le \lambda].$$

(2) *If* $\lim_{n\to\infty} \frac{1}{n}$ $[\#$ *eigenvalues of* $S_n < \lambda]$ *exists for all* λ *in* $\mathbb{R}$, *then*

$$\tfrac{1}{2}\log\left[\frac{B+1}{\lambda+1}\right]dF(\lambda) \le C_W^{\infty}(P) \le \tfrac{1}{2}\int_{-1}^{B} \log\left[\frac{B+1}{\lambda+1}\right]dF(\lambda) + \overline{\lim_{n}}\ \frac{1}{n} H_{GN}^{n}(N)$$

where F is a distribution function defined by

$$F(\lambda) = \lim_{n\to\infty} \frac{1}{n}[\#\ \text{eigenvalues of}\ S_n \le \lambda] = \lim_{n\to\infty} F_n(\lambda),$$

and the constant B is defined by

$$P = \int_{-1}^{B}[B-\lambda]dF(\lambda).$$

(3) *If* $\overline{\lim_{n}}\ \frac{1}{n} H_{GN}^{n}(N) = 0$, *then* $C_W^{\infty}(P) = 0$ *if and only if* $\overline{\lim_{n\to\infty}}\ \frac{1}{n}$ $[\#$ *eigenvalues of* $S_n < \lambda] = 0$ *for all* λ *in* $\mathbb{R}$. *This requires that S be unbounded and always occurs if* $+\infty$ *is the only limit point of* $\sigma(S)$.

<u>Proof</u>. It is sufficient to evaluate $\overline{\lim_{n}}\ \frac{1}{n} C_W^{n}(nP)$. Applying Lemma 3 to the expressions contained in Lemma 1,

$$\frac{1}{2}\int_{-1}^{B_n} \log\left[\frac{B_n+1}{\lambda+1}\right]dG_n(\lambda) \le \frac{1}{n} C_W^n(nP) \le \frac{1}{2}\int_{-1}^{B_n} \log\left[\frac{B_n+1}{\lambda+1}\right]dG_n(\lambda) + \frac{1}{n} H_{GN}^n(N)$$

where $G_n(\lambda) = \frac{1}{n}$ [# eigenvalues of $S_n < \lambda$] and $P = \int_{-1}^{B_n} [B_n-\lambda]dF_n(\lambda)$, $n \ge 1$.

Note that since the integrands appearing in the statement of (1) are equal to zero when evaluated at the upper limit, and are continuous, the left-continuous function $G_n(\lambda) = \frac{1}{n}$ [# eigenvalues of $S_n < \lambda$] can be replaced by $G_n(\lambda^+) = F_n(\lambda)$. The same comment holds for the remainder of the theorem.

To prove (2), we can suppose that $\overline{\lim_n} \frac{1}{n} H_{GN}^n(N) = 0$. Then, let (n_j) be any subsequence of the integers such that

$$\overline{\lim_n} \frac{1}{n} C_W^n(nP) = \lim_{j\to\infty} \frac{1}{2} \int_{-1}^{B(n_j)} \log\left[\frac{B(n_j)+1}{\lambda+1}\right]dF_{n_j}(\lambda).$$

Let B be any finite limit point of $(B(n_j),\ j \ge 1)$ (assuming that at least one exists). Then, restricting attention to (n_j) such that $B(n_j) \to B$,

$$2C_W^\infty(P) = \lim_{j\to\infty} \int_{-1}^{B(n_j)} \log\left[\frac{B(n_j)+1}{\lambda+1}\right]dF_{n_j}(\lambda)$$

$$= \lim_{j\to\infty}\left[\int_{-1}^{B} \log\left[\frac{B(n_j)+1}{\lambda+1}\right]dF_{n_j}(\lambda) - \int_{B(n_j)}^{B} \log\left[\frac{B(n_j)+1}{\lambda+1}\right]dF_{n_j}(\lambda)\right].$$

Now

$$\lim_{j\to\infty}\left|\int_{B(n_j)}^{B} \log\left[\frac{B(n_j)+1}{\lambda+1}\right]dF_{n_j}(\lambda)\right| \le \overline{\lim_{j\to\infty}}\left|\log \frac{B(n_j)+1}{B+1}\right|\left|(F_{n_j}(B) - F_{n_j}(B(n_j))\right|$$

$$\le \lim_{j\to\infty} \left|\log\left[\frac{B(n_j)+1}{B+1}\right]\right| = 0.$$

We thus obtain

$$2C_W^\infty(P) = \lim_{j\to\infty} \int_{-1}^{B} \log\left[\frac{B(n_j)+1}{\lambda+1}\right] dF_{n_j}(\lambda)$$

$$= \lim_{j\to\infty} \int_{-1}^{B} \left[\log[\frac{B(n_j)+1}{B+1}] + \log[\frac{B+1}{\lambda+1}]\right] dF_{n_j}(\lambda)$$

and since $\left|\int_{-1}^{B} \log\left[\frac{B(n_j)+1}{B+1}\right] dF_{n_j}(\lambda)\right| \le \left|\log\left[\frac{B(n_j)+1}{B+1}\right]\right|$, one has

$$C_W^\infty(P) = \lim_{j\to\infty} \tfrac{1}{2} \int_{-1}^{B} \log[\frac{B+1}{\lambda+1}] dF_{n_j}(\lambda)$$

and, similarly,

$$P = \lim_{j\to\infty} \int_{-1}^{B} [B - \lambda] dF_{n_j}(\lambda).$$

Assuming that $F_{n_j}(\lambda) \to F(\lambda)$ for all λ in $\mathbb{R}$, this gives (2) of the Theorem whenever $\{B(n_j), j \ge 1\}$ has a finite limit point. Moreover, in this case it can be seen that every maximizing sequence must have B as a limit point.

To complete the proof of (2) and simultaneously prove (3), suppose that (n_j) is as above and that no maximizing sequence $(B(n_j), j > 1)$ has a finite limit point. Since $P = \int_{-1}^{B(n_j)} [B(n_j) - \lambda] dF_{n_j}(\lambda)$ for all $j > 1$, and $B(n_j) \to \infty$, it follows that $\lim_{j\to\infty} F_{n_j}(\lambda) = F(\lambda) = 0$ for every finite λ and every maximizing sequence (n_j). Thus, in this case $C_W^\infty(P)$ has the same value as if $+\infty$ were the only limit point of $\sigma(S)$; assume WLOG that this is true. For a fixed n, the channel in H_n is given by $X + \tilde{P}_n N = Y$, where $\mu_X[H_n] = 1$ and $E_{\mu_X}\|x\|^2_{W,n} \le nP$. A given x in H_n has the representation $x = V_n R_{W,n}^{1/2} v$ for some unique v in H_n, and $\|x\|_{W,n} = \|v\|$. V_n is isometric on range$(R_W^{1/2}\tilde{P}_n)$. The H_n channel is thus equivalent to the H_{W_n} channel $Y = X + V_n^{*}\tilde{P}_n N$, with X a.s. in $H_{W,n}$, $E_{\mu_X}\|x\|^2_W \equiv E_{\mu_X}\|R_W^{-1/2}x\|^2 \le nP$. For this n-dimensional channel, $C_W^n(nP) \le C_W^n(nP, \max)$, where

$C_W^n(nP, \max)$ is the information capacity of the infinite-dimensional Gaussian channel when the noise has covariance operator R_N and the transmitted signal measure μ_X must satisfy: $\dim[\mathrm{supp}(\mu_X)] \leq n$, $\mu_X[\mathrm{range}(R_W^{1/2})] = 1$, and $E_{\mu_X} \|x\|_W^2 \leq nP$. From Theorem 2 of [2]

$$C_W^n(nP, \max) = \tfrac{1}{2} \sum_{i=1}^{K(n)} \log\left[\frac{C(n)}{\lambda_i}\right]$$

where $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots$ are the eigenvalues of I+S (since S has a finite set of limit points, it has pure point spectrum);

$$P = \frac{1}{n} \sum_{i=1}^{K(n)} [C(n) - \lambda_i]$$

and

$$K(n) = \sup\{i: \lambda_i < C(n)\}.$$

Thus,

$$\frac{1}{n} C_W^n(nP, \max) \leq \frac{1}{2n} \sum_{i=1}^{K(n)} \left[\frac{C(n)-\lambda_i}{\lambda_i}\right]$$

and since $\lambda_i \nearrow \infty$ and $\frac{1}{2n} \sum_{i=1}^{K(n)} [C(n)-\lambda_i] = \frac{P}{2}$, this gives

$$\overline{\lim_n} \frac{1}{n} C_W^n(nP) \leq \overline{\lim_n} \frac{1}{n} C_W^n(nP, \max) = 0.$$

Thus, for a maximizing subsequence (n_j), either $(B(n_j))$ has a finite limit point B, or else $B(n_j) \to \infty$. In the first case, (1) and (2) (when applicable) of the Theorem give the capacity. In the second case, $C_W^\infty(P) = 0$. It is clear that either every maximizing sequence (n_j) has $(B(n_j))$ with a finite limit point, which has the unique value B, or else $(B(n_j))$ has no finite limit point for any maximizing sequence (n_j). □

Remark. In part (2), the same result holds so long as $F(\lambda) \equiv \lim_{n\to\infty} F_n(\lambda)$ exists for all $\lambda < B$, with B defined as in (2).

The following result is a slight refinement of part (1) of the Theorem; the direct proof is omitted.

Corollary 1. Suppose that $\overline{\lim_n} \frac{1}{n} C_W^n(nP) > 0$ and that $\overline{\lim_n} \frac{1}{n} H_{GN}^n(N) = 0$, and let B be the largest number such that $P \geq \overline{\lim_n} \int_{-1}^{B} [B_n - \lambda] dF_n(\lambda)$. Then

$$\overline{\lim_n} \frac{1}{n} C_W^n(nP) = \overline{\lim_n} \int_{-1}^{B} \log\left[\frac{B_n+1}{\lambda+1}\right] dF_n(\lambda),$$

where (B_n) and (F_n) are defined as in Theorem.

Corollary 2. When $\overline{\lim_n} \frac{1}{n} H_{GN}^n(N) = 0$, then bounds on $C_W^\infty(P)$ are given by

$$\tfrac{1}{2} \log(1 + P/\lambda_{max}) \leq C_W^\infty(P) \leq \tfrac{1}{2} \log(1 + P/\lambda_{min})$$

where λ_{min} is the smallest limit point in the spectrum of S, and λ_{max} is the largest. Moreover, these bounds can be attained by proper choice of (H_{W_n}).

Proof. The bounds follow from Prop. 2. The upper bound is attained if $\{u_i, i \geq 1\}$ is chosen so that $\|(S-\gamma I)u_i\| \to \lambda_{min}$; such an o.n. set exists, since λ_{min} is in $\sigma_{ess}(S)$. The lower bound is attained in a similar manner. □

Applications

A number of applications of the main result are given in [3]. These include new results for nonstationary discrete-time channels and for both stationary and non-stationary continuous-time channels when the time of transmission is fixed.

Extensions

For the classical continuous-time channel, the time of transmission is permitted to increase in order to define average information capacity. The dimensionality n used here is replaced by the transmission time T, and the subspace H_n is replaced by $L_2[0,T]$. The constraint $E_{\mu_X^n}\|x\|^2_{W,n} \leq nP$ is replaced by the constraint $E_{\mu_X^T}\|x\|^2_{W,T} \leq TP$, where $\mu_X^T[\text{range}(R^{1/2}_{W,T})] = 1$, $\|x\|^2_{W,T} \equiv \|R^{-1/2}_{W,T}x\|_T$, and $\|\cdot\|_T$ is the $L_2[0,T]$ norm. $R_{W,T}$ is a bounded linear operator in $L_2[0,T]$ defined by a covariance function r_W, with $r_W(t,s)$ defined for all s, $t \geq 0$.

The method used here can also be used to analyze capacity of such channels. Those results will be given elsewhere.

References

[1] R.B. Ash, *Information Theory*, Interscience, New York, 1965.

[2] C.R. Baker, Capacity of the mismatched Gaussian channel, *IEEE Trans. on Inform. Theory*, IT-33, 802-812 (1987).

[3] C.R. Baker, Coding capacity of a class of additive channels, to appear.

[4] R.G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York (1968).

[5] S. Ihara, Capacity of discrete-time Gaussian channel with and without feedback, I, *Memoirs of the Faculty of Science*, Kochi Univ., Series A, Mathematics, 9, 21-36 (1988).

[6] S. Ihara, On the capacity of channels with additive nonGaussian noise, *Inform. Control*, 37, 34-39 (1978).

[7] F. Riesz and B. Sz-Nagy, *Functional Analysis*, Ungar, New York, 1955.